

SIHONG XIE

Phone: (773) 681-1984 Email: xiesihong1@gmail.com

Add: Room 326, Mountaintop Campus, Lehigh University, Bethlehem, PA 18018

EDUCATION

University of Illinois at Chicago

PhD in Computer Science

Sep, 2010 - Aug, 2016

Sun Yat-Sen University, Guangzhou, China

M.E. in Software Engineering

Sep, 2008 - Jun, 2010

B.E. in Software Engineering

Sep, 2004 - Jun, 2008

EMPLOYMENT

Lehigh University, Bethlehem, PA

Aug, 2016 - *present*

Assistant Professor, Department of Computer Science and Engineering

University of Illinois at Chicago, Chicago, IL

Aug, 2012 - Aug, 2016

Research Assistant, Department of Computer Science

Supervisor: Philip S. Yu

LinkedIn Corporation, Mountain View, CA

May, 2013 - Aug, 2013

Applied Researcher, Search and Network Analysis Group

Supervisors: Anmol Bhasin, Baoshi Yan, Mohammad Amin

IBM Research, Hawthorne, NY

May, 2012 - Aug, 2012

Research Intern, Exploratory Stream Analytics Group

Supervisor: Wei Fan

RESEARCH FUNDING

- **NSF**, “*CAREER: Bilevel Optimization for Accountable Machine Learning on Graphs*”. \$556,481. 2022-2027
- **Lehigh Research Future**, “*Mathematical Optimization and Data Science*”. \$250,000. **Co-PI**. 2021
- **NSF**, “*III: Small: Collaborative Research: Algorithms, systems, and theories for exploiting data dependencies in crowdsourcing*”. \$500,000. Lehigh: \$250,000. **Lead PI**. 2020-present
- **NSF**, “*SaTC: CORE: Small: Collaborative: Learning Dynamic and Robust Defenses Against Co-Adaptive Spammers*”, NSF. \$500,000. Lehigh: \$250,000. **Co-PI**. 2019-present
- **Lehigh Accelerator** “*Efficient, Explainable and Robust Data Scientific Methods for Smart Engineering Systems*”. \$100,000. **Co-PI**. 2019-present

Under submission/preparation

- [1] Yazheng, Liu, Xi, Zhang, and **Sihong Xie** Explaining GNN over Evolving Graphs using Information Flow. Under submission.
- [2] Yuefei, Liu, Xi, Zhang, and **Sihong Xie** Interpretable reinforcement attacks against graph-based rumor detectors. Under submission.
- [3] Jiaxin Liu and **Sihong Xie**. Subgroup discovery for fair node classification on graphs. Under submission.
- [4] Chao Chen, Chenghua Guo, Xi Zhang, and **Sihong Xie**. Robust explanations of Siamese networks. Under preparation.
- [5] Han Yu, Jiaxin Liu, Chao Chen, Qi Li, Xi Zhang, and **Sihong Xie**. Explainable uncertainty quantification of graphical models using linear bound propagation. Under preparation.

Refereed Conferences

- [1] Yifei Liu, Chao Chen, Yazheng Liu, Xi Zhang, and **Sihong Xie**. Multi-objective Explanations of GNN Predictions. ICDM, 2021. (*Regular paper, acceptance rate 10%. Selected for publication on the journal KAIS*).
- [2] Chao Chen, Yifan Shen, Guixiang Ma, Xiangnan Kong, Srinivas Rangarajan, Xi Zhang, and **Sihong Xie**. Self-learn to Explain Siamese Networks Robustly ICDM, 2021. (*Short paper, acceptance rate 20%*).
- [3] Nasim Sabetpour, Adithya Kulkarni, **Sihong Xie**, and Qi Li. Truth Discovery in Sequence Labels from Crowds ICDM, 2021. (*Regular paper, acceptance rate 10%. Selected for publication on the journal KAIS*).
- [4] Kai Burkholder, Kenny Kwok, Yuesheng Xu, Jiaxin Liu, Chao Chen, and **Sihong Xie**. Certification and Trade-off of Multiple Fairness Criteria in Graph-based Spam Detection. CIKM, 2021. (*Full paper, acceptance rate 21.7%*).
- [5] Hang Yin, John Lee, Thomas Hartvigsen, Xiangnan Kong, and **Sihong Xie**. Energy-Efficient Models for High-Dimensional Spike Train Classification using Sparse Spiking Neural Network. KDD, 2021. (*Full paper, acceptance rate 15.5%*).
- [6] Shengli Zhu, Jakob Coles, and **Sihong Xie**. Active Search using Meta-Bandits. CIKM, 2020. (*Poster paper*).
- [7] Yifei Liu, Chao Chen, Yazheng Liu, Xi Zhang, and **Sihong Xie**. Shapley Values and Meta-Explanations for Probabilistic Graphical Model Inference. CIKM, 2020. (*Full paper, acceptance rate 20%*).
- [8] Yingtong Dou, Guixiang Ma, Philip S. Yu, and **Sihong Xie**. Robust Spammer Detection by Nash Reinforcement Learning. KDD, 2020. (*Full paper, acceptance rate 16.9%*).

- [9] Chao Chen, Yifei Liu, Xi Zhang, and **Sihong Xie**. Scalable Explanation of Inferences on Large Graphs. ICDM, 2019. (*Short paper, acceptance rate 18.5%*).
- [10] Youshan Zhang, **Sihong Xie**, and Brian D. Davison. Transductive Learning via Improved Geodesic Sampling. BMVC, 2019. (Regular paper, acceptance rate 28%).
- [11] Shuaijun Ge, Guixiang Ma, **Sihong Xie**, and Philip S. Yu. Securing Behavior-based Opinion Spam Detection. BigData, 2018. (Regular paper, acceptance rate 19.7%).
- [12] Lei Zheng, Yixue Wang, Lifang He, **Sihong Xie**, Fengjiao Wang, and Philip S. Yu. PER: A Probabilistic Attentional Model for Personalized Text Recommendations. BigData, 2018. (Regular paper, acceptance rate 19.7%).
- [13] Vahid Noroozi, Sara Bahaadini, Lei Zheng, **Sihong Xie**, Weixiang Shao, and Philip S. Yu. Semi-supervised Deep Representation Learning for Multi-View Problems. BigData, 2018. (Regular paper, acceptance rate 19.7%).
- [14] **Sihong Xie** and Philip S. Yu. Next Generation Trustworthy Fraud Detection. CIC, 2018.
- [15] Hu Xu, **Sihong Xie**, Shu Lei, and Philip S. Yu. Dual attention network for product compatibility and function satisfiability analysis. AAAI, 2018. (Regular paper, acceptance rate 25.8%).
- [16] Hu Xu, **Sihong Xie**, Lei Shu, and Philip S. Yu. Product function need recognition via semi-supervised attention network. BigData, 2017. (Regular paper, acceptance rate 17.8%).
- [17] Xiaokai Wei, **Sihong Xie**, Bokai Cao, and Philip S. Yu. *Rethinking Unsupervised Feature Selection: From Pseudo Labels to Pseudo Must-Links*. ECML. 2017. (Regular paper, acceptance rate 29%).
- [18] Vahid Noroozi, Lei Zheng, Sara Bahaadini, **Sihong Xie**, and Philip S. Yu. Seven: Deep Semi-supervised Verification Networks. IJCAI, 2017. (Regular paper, acceptance rate 25.9%).

Journals

- [19] Xi Zhang, Yuan Su, Siyu Qu, **Sihong Xie**, Binxing Fang and Philip S. Yu. IAD: Interaction-Aware Diffusion Framework in Social Networks *IEEE Transactions on Knowledge and Data Engineering (TKDE)*, 2018. Impact Factor: 3.857.
- [20] **Sihong Xie** and Philip S. Yu. Active zero-shot learning: A novel approach to extreme multi-labeled classification. *International Journal of Data Science and Analytics*, 2016.

Before joining Lehigh

Journal

- [21] Senzhang Wang, **Sihong Xie**, Xiaoming Zhang, Zhoujun Li, Philip S. Yu, and Yueying He. Coranking the future influence of multiobjects in bibliographic network through mutual reinforcement. *ACM Trans. Intell. Syst. Technol.*, 7(4):64:1–64:28, May 2016.
- [22] Guan Wang, **Sihong Xie**, Bing Liu, and Philip S. Yu. Identify online store review spammers via social review graph. *ACM Trans. Intell. Syst. Technol.*, 3(4):61:1–61:21, September 2012.

Refereed Conference

- [23] **Sihong Xie**, Qingbo Hu, Weixiang Shao, Jingyuan Zhang, Jing Gao, Wei Fan, and Philip S. Yu. Effective crowd expertise modeling via cross domain sparsity and uncertainty reduction. *SDM*. SIAM, 2016.
- [24] **Sihong Xie**, Shaoxiong Wang, and Philip S. Yu. Active zero-shot learning. *CIKM*, 2016.
- [25] Chun-Ta Lu, **Sihong Xie**, Weixiang Shao, Lifang He, and Philip S. Yu. Item recommendation for emerging online businesses. *IJCAI*, 2016.
- [26] Yuan Yuan, **Sihong Xie**, ChunTa Lu, Jie Tang, and Philip S. Yu. Interpretable and effective opinion spam detection via temporal patterns mining across websites. *BigData*. IEEE, 2016.
- [27] Chenwei Zhang, **Sihong Xie**, Yaliang Liu, Jing Gao, Wei Fan, and Philip S. Yu. Multi-source hierarchical prediction consolidation. *CIKM*, 2016.
- [28] J. Zhang, C. T. Lu, M. Zhou, S. Xie, Y. Chang, and P. S. Yu. Heer: Heterogeneous graph embedding for emerging relation detection from news. In *2016 IEEE International Conference on Big Data (Big Data)*, 2016.
- [29] Jingyuan Zhang, Bokai Cao, **Sihong Xie**, Chun-Ta Lu, Philip S. Yu, and Ann B. Ragin. Identifying connectivity patterns for brain diseases via multi-side-view guided deep architectures. *SDM*. SIAM, 2016.
- [30] Qingbo Hu, **Sihong Xie**, Jiawei Zhang, Qiang Zhu, Songtao Guo, and Philip S. Yu. Heterosales: Utilizing heterogeneous social networks to identify the next enterprise customer. *WWW*. ACM, 2016.
- [31] Jingyuan Zhang, Luo Jie, Altaf Rahman, **Sihong Xie**, Yi Chang, and Philip S. Yu. Learning entity types from query logs via graph-based modeling. *CIKM*. ACM, 2015.
- [32] Xiaokai Wei, **Sihong Xie**, and Philip S. Yu. Efficient partial order-preserving unsupervised feature selection on networks. In *SDM*. SIAM, 2015.
- [33] **Sihong Xie**, Qingbo Hu, Jingyuan Zhang, Jing Gao, Wei Fan, and Philip S. Yu. Robust crowd bias correction via dual knowledge transfer from multiple overlapping sources. In *BigData*. IEEE, 2015.
- [34] **Sihong Xie**, Qingbo Hu, Jingyuan Zhang, and Philip S. Yu. An effective and economic bi-level approach to ranking and rating spam detection. In *DSAA*. IEEE, 2015.
- [35] **Sihong Xie**, Jing Wang, Mohammad S.Amin, Baoshi Yan, Anmol Bhasin, Clement Yu, and Philip S. Yu. A context-aware approach to detection of short irrelevant texts. In *DSAA*. IEEE, 2015.
- [36] Bowen Dong, **Sihong Xie**, Jing Gao, Wei Fan, and Philip S. Yu. Onlinecm: online consensus maximization with missing values. In *SDM*. SIAM, 2015.
- [37] Qingbo Hu, **Sihong Xie**, Shuyang Lin, Senzhang Wang, and Philip S. Yu. Clustering embedded approaches for efficient information network inference. *Data Science and Engineering*, 2015.

- [38] Qiongbo Hu, **Sihong Xie**, Shuyang Lin, Wei Fan, and Philip S. Yu. Frameworks to encode user preferences for inferring topic-sensitive information networks. In *SDM*. SIAM, 2015.
- [39] Qiongbo Hu, **Sihong Xie**, Shuyang Lin, Senzhang Wang, and Philip S. Yu. CENI: a hybrid framework for efficiently inferring information networks. In *ICWSM*. AAAI, 2015.
- [40] **Sihong Xie**, Jing Gao, Deepak Turaga, Wei Fan, and Philip S. Yu. Class-distribution regularized consensus maximization for alleviating overfitting in model combination. In *KDD*. ACM, 2014.
- [41] Senzhang Wang, **Sihong Xie**, Xiaoming Zhang, Zhoujun Li, Philip S. Yu, and Xinyu Shu. Future influence ranking of scientific literature. In *SDM*. SIAM, 2014.
- [42] Guoqiong Liao, Philip S. Yu, Qianhui Zhong, **Sihong Xie**, Zhen Shen, Changxuan Wan, and Dexi Liu. Trajectory event cleaning for mobile rfid objects. In *15th International Conference on Mobile Data Management*. IEEE, 2014.
- [43] Guoqiong Liao, Yuchen Zhao, **Sihong Xie**, and Philip S. Yu. An effective latent networks fusion based model for event recommendation in offline ephemeral social networks. In *CIKM*. ACM, 2013.
- [44] Chun-Ta Lu, **Sihong Xie**, Xiangnan Kong, and Philip S. Yu. Inferring the impacts of social media on crowdfunding. In *WSDM*. ACM, 2013.
- [45] **Sihong Xie**, Xiangnan Kong, Jing Gao, Wei Fan, and Philip S. Yu. Multilabel consensus classification. In *ICDM*. IEEE, 2013.
- [46] Chuan Shi, Xiangnan Kong, Philip S. Yu, **Sihong Xie**, and Bin Wu. Relevance search in heterogeneous networks. In *EDBT*. ACM, 2012.
- [47] **Sihong Xie**, Wei Fan, and Philip S. Yu. An iterative and re-weighting framework for rejection and uncertainty resolution in crowdsourcing. In *SDM*. SIAM, 2012.
- [48] **Sihong Xie**, Guan Wang, Shuyang Lin, and Philip S. Yu. Review spam detection via temporal pattern discovery. In *KDD*. ACM, 2012.
- [49] **Sihong Xie**, Guan Wang, Shuyang Lin, and Philip S. Yu. Review spam detection via time series pattern discovery. In *WWW (short paper)*. ACM, 2012.
- [50] Guan Wang, **Sihong Xie**, Bing Liu, and Philip S. Yu. Review graph based online store review spammer detection. In *ICDM*. IEEE, 2011.
- [51] **Sihong Xie**, Wei Fan, Olivier Verscheure, and Jiangtao Ren. Efficient and numerically stable sparse learning. In *ECML/PKDD*, 2010.
- [52] **Sihong Xie**, Wei Fan, Jing Peng, Olivier Verscheure, and Jiangtao Ren. Latent space domain transfer between high dimensional overlapping distributions. In *WWW*. ACM, 2009.
- [53] Erheng Zhong, Sihong Xie, Wei Fan, Jiangtao Ren, Jing Peng, and Kun Zhang. Graph-based iterative hybrid feature selection. In *ICDM*. IEEE, 2008.

WORKSHOP ORGANIZING

The Third International Workshop on Truth Discovery and Fact Checking: Making a Credible Web for Tomorrow, Virtual Workshop co-host with KDD 2021.

The Second International Workshop on Truth Discovery and Fact Checking: Making a Credible Web for Tomorrow, Virtual Workshop co-host with KDD 2020.

INVITED TALKS AND LECTURES

Fairness in machine learning

Outreach to K6-12 students at LEAP Academy, Rutgers University, Camden Nov, 2021

Robust misinformation detection for e-commerce.

Seminar, Meta (former Facebook) Nov, 2021

Interpretable, Robust, and Fair Learning on Graphs.

Seminar, CS Department at Iowa State University Oct, 2021

Seminar, OptML group at Lehigh ISE department Mar, 2021

Secure Opinion Spam Detection

The Second International Workshop on Truth Discovery and Fact Checking Aug, 2020

MARS: Memory Attention-Aware Recommender System

DSAA 2019, Washington, DC Sep, 2019

Securing Behavior-based Opinion Spam Detection

BigData 2018, Seattle, WA Dec, 2018

Next Generation Trustworthy Fraud Detection

CIC 2018, Philadelphia, PA Oct, 2018

WPI, MA Dec, 2018

Crowdsourcing for Natural Language Processing applications.

Beijing University of Posts and Telecommunications, Beijing, China. Jun, 2017

Sun Yat-Sen University, Guangzhou, China. Jul, 2017

Exploiting multiple sources for trustworthy content managements

BigData Conference 2015, Santa Clara, CA Nov, 2015

Enhancing predictive power via consensus models

University of Illinois, Chicago, IL Nov, 2014

Overfitting and regularization in model combination

KDD 2014, New York City, NY Aug, 2014

Multi-label consensus classification

ICDM 2013, Dallas, TX Dec, 2013

Context-aware spam detection

LinkedIn Corporation, Mountain View, CA Aug, 2013

HONORS AND AWARDS

Faculty Grants for International Connections, Lehigh University,	2017
SIAM SDM 2012 Student Travel Award	May, 2012
ACM SIGKDD 2012 Student Travel Award	Aug, 2012
IEEE ICDM 2008 Data Mining Contest Crown Award (highest award)	2008
Outstanding Graduate Student Scholarship, Sun Yat-Sen University	2008 - 2009
Outstanding Student Scholarship First Prize, Sun Yat-Sen University	2006 - 2007

TEACHING

15 offerings of 4 courses from 2016 Fall - 2022 Spring.

CSE 326/426 “Fundamentals of Machine Learning”, Spring 2019, Fall 2020, Fall 2021.

CSE 325/425 “Natural Language Processing”, Fall 2017, Fall 2018, Fall 2019, Spring 2021, Spring 2022.

CSE 241 “Database Systems & Apps”, Spring 2017, Spring 2018, Fall 2019, Spring 2020, Fall 2020, Fall 2021.

CSE 398/498 “Text Mining”, Fall 2016

CSE 406 “Research Methods”, Guest Lecturer, Fall 2016-2017, 2018, 2020

RESEARCH ADVISING

Doctorals

Chao Chen, CSE. 2018 Summer - present.

Jiixin Liu, CSE. 2019 Fall - present.

Xuehan Chen, CSE. 2021 Fall - present.

Masters

Praveer Narwelkar, CSE. 2021 Spring - Fall . *Multi-modal misinformation detection.*

Yue Shu, ECE. 2018 Spring. *Multiple instance learning for sentimental information extraction.*

Kexin Ding, ECE. 2018 Spring. *Building a web interface for crowdsourcing a multi-modal (text-image) sentimental dataset.*

Undergraduates

Sherry Huang, Lehigh University. 2022 Spring - present. *Fair machine learning.*

Eric Enouen, Ohio State University. 2021 Summer - present. *Constrained multi-objective optimization.*

Kenny Kwock, Lehigh University. 2020 Summer - 2021 Fall. *Fair multi-objective spam detection.*

Kieran O’Connor, Lehigh University. 2021 Spring - Fall. *Explainable graphical models.*

Kai Burkholder, Lehigh University. 2020 Summer. *Fair multi-objective spam detection.*

Basilio Garcia Castillo, CSE. 2016 Fall - 2017 Spring. *Rule-based sentiment analysis.* Now working for Google NYC

Tom Bousso, CS at UC Berkeley, CA. 2017 summer. *Active learning for implicit sentiment analysis.*

Hannah Lambert, Chijioke Umezina, Jun-Lucas Pritsker, CSE. 2017 Fall *Senior Design: crowd-sourcing for aspect-based sentiment analysis*.

Zhenyu Li, ISE. 2017 Fall. *Fine-grained sentimental information extraction using topic models*.

Yilin Zhu, CS at De Anza College, Santa Clara, CA. 2017 Summer. *Spam detection using graphical model*.

Jun-Lucas Pritsker, CSE. 2017 Winter - 2018 Spring. *Human-like sentimental mapping from images to texts*. co-adviser: Xiaolei Huang.

Kexin Shu, CS at De Anza College, Santa Clara, CA. 2019 Summer. *Spam detection*.

Jakob Coles, Lehigh University. 2019 Summer - 2020 Spring (REU). *Graphical bandits*.

Jacob Nocentino. 2019 Summer (REU) . *Graphical bandits*.

Jeremy Ding, Lehigh University. 2019 Summer. *Graphical bandits*.

Sheldon Xu, Lehigh University. 2019 Summer - 2021 Spring. *Graphical bandits, fair spam detection*.

Sam Chebruch, Lehigh University. 2019 Summer - 2020 Spring. *Graphical bandits*.

Shengli Zhu, Lehigh University. 2017 Summer - 2020 Spring. *Graphical bandits, relation extraction, spam detection game*.

Yifan Zhang, Lehigh University. 2020 Summer - 2020 Fall. *Explainable spam detection*.

Dissertation committee

Weihang Yuan, CSE, Lehigh University. 2022 Spring. General exam. *Deep Goal Reasoning*.

Morgan Fine-Morris, CSE, Lehigh University. 2022 Spring. General exam. *Learning Planning Knowledge in Structured Domains with Numeric Fluents*.

Noah Reifsnnyder, CSE, Lehigh University. 2022 Spring. Defense. *A Study of Goal Reasoning Agents Expectations for Nondeterministic and Continuous domains*.

Suyun Liu, ISE, Lehigh University. 2021 Fall. General exam. *Stochastic Multi-Objective Optimization and Its Application to Fairness in Machine Learning*.

Sicong Kuang, CSE, Lehigh University. 2020 Spring. Defense. *Improving Semantic Representation in Bias Detection*.

Zhiyu Chen, CSE, Lehigh University. 2021 Spring. General exam. *Dataset Search and Augmentation*.

SERVICES

Conference and Journal Reviewer

Senior PC member: AAAI (2022).

PC member: ICLR (2022), SDM (2017-2022), ICDM (2017-2021), KDD (2018-2022), CIKM (2017-2021), AAAI (2018-2022), IJCAI (2020), SIGIR (2020-2022).

External conferece reviewer: SIGAPP SAC (2012), APWeb (2012), PRICAI (2012), ICDM (2012 - 2014), ASONAM (2013), WWW (2014), SDM (2014 - 2016), BigData (2015), AAAI (2016).

Journal reviewer: the Journal of AI Research (JAIR) (2021), Frontiers in Big Data (2022), Transactions on Knowledge Discovery from Data (TKDD) (2014, 2016 - 2018), Transactions on Knowledge and Data Engineering (TKDE) (2017-2020, 2022), and the Transactions on the Web (TWeb) (2018, 2019, 2022).

Conference session chairs

SDM (2021), SIGIR (2021), ICDM (2021)

Grant Panelist

NSF Panelist (2017, 2020, 2021, 2022)

University

CSE department: DE&I committee. 2021-2022

CSE department: Faculty hiring committee. 2020-2021

University: Faculty Senate (Faculty affair sub-committee). 2020 Fall

CSE department: Curriculum committee. 2016 - 2020

CSE department: Graduate Admissions Committee. 2016 - 2020